A DIETARY ANALYSIS OF THE MICROWHIP SCORPION (ARACHNIDA: PALPIGRADI) FROM VAL VERDE COUNTY, TEXAS

A Thesis

Presented to the

Faculty of the College of Graduate Studies and Research

Angelo State University

In Partial Fulfilment of the

Requirements for the Degree

MASTER OF SCIENCE

by

MARY PATRICIA JONES

May 2017

Major: Biology

A DIETARY ANALYSIS OF THE MICROWHIP SCORPION (ARACHNIDA: PALPIGRADI) FROM VAL VERDE COUNTY, TEXAS

by

MARY PATRICIA JONES

APPROVED:

Dr. Ned Strenth

Dr. Loren Ammerman

Dr. Nicholas Negovetich

Dr. Andrew Siefker

4 April 2017

APPROVED:

Dr. Susan Keith
Dean, College of Graduate Studies and Research

DEDICATION

I would like to dedicate this effort with gratitude to my grandparents, G. Emerson and Patricia Jones and Frederick and Mary Rock. You laid the foundation for a love of learning and hard work; you have inspired many.

ACKNOWLEDGEMENTS

I would like to first and foremost thank my advisor and mentor, Dr. Ned Strenth. Thank you for taking the time to first get me involved in research and for your continued guidance ever since. I truly value your insight, experience, and passion for teaching. I hope to continue to use the lessons you have taught me to become a better biologist, teacher, and researcher. In addition, I would like to thank Mrs. Barbara Strenth, for her enduring patience.

Thank you to Mr. Lynn McCutchen, for arachnid expertise, for first finding the BORR palpigrades, showing me how to find them, and for your impressive photographs.

I am very grateful to Dr. Bob and Mrs. Jean Ann LeGrand for generous access to the Big Oak River Ranch (BORR).

Thank you to my thesis committee; Dr. Loren Ammerman, Dr. Nicholas Negovetich, and Dr. Andrew Siefker. I appreciate the patient cooperation, input, and guidance you have provided during this project.

I am exceedingly thankful to Angelo State University for the numerous opportunities you have provided. To the Biology Department faculty, staff, and students for the many lessons you have taught me and your continual support. A special thanks to Mrs. Jody Casares who makes it all function smoothly. Thank you to the office of the graduate dean, Dr. Susan Keith, Jana Sparks, and Katie Ede, for always being willing to work with me and answering countless questions. Mr. Brian Beck from the ASU IT department for your enthusiasm and timely assistance. Additionally to Erin Adams and Mrs. Ewa Davis for advice and encouragement.

For funding this project, I extend gratitude to the Head of the River Ranch (HORR) Graduate Research Scholarship, established by Frank and Edith Boulware and currently

maintained by Mr. Rylan Howard. Also, the Angelo State University Office of Sponsored Projects for awarding the Graduate Research Fellowship and to Ms. Katie Marie Plum and Ms. Elizabeth Randall for your dedication to students.

I am very appreciative to Dr. John Hanson, and the staff at the Research and Testing Genomics Lab as well as Dr. Chris Bird, Mrs. Sharon Magnuson, and especially Evan Krell, from the TAMUCC, Core Genomics Lab for your expertise and facilities.

I would also like to thank many volunteers who helped me search for, (and on the lucky occasion found), these elusive palpigrades; Griffin Chodacki, Travis Fisher, Paul Ford, Haley Hale, Brandy Hall, Casey Harrison, Chance Hoff, P. Citlally Jimenez, Kerri Kelley, Ciera Lanier, Michael Lucero, Timothy Maddox, Annabelle Mellor, Dr. Nicholas Negovetich, Austin Osmanski, Joey Prien, Kalin Skinner, Dr. Benjamin Skipper, Dustin Tarrant, Gizell Vasquez, Jacob Wright, and Tom Zimmerman.

Many thanks to the Holik's; Marilyn and Benny, Gayla and Adolph, Becky and Ronnie, Brenda and "Sunny-boy", John and Kim, and your families for all your kindness and care.

And finally to my family, George, Kristine, Rachel, Caroline, and Rebecca Jones for your constant love, ceaseless encouragement, and unfailing support.

ABSTRACT

Palpigrades comprise a poorly known group of arachnids found in caves and soil from numerous localities worldwide. Prior and preliminary studies have suggested arthropods and cyanobacteria as possible diet items of these organisms. This current study uses DNA sequencing to identify contents of the digestive tract of eleven palpigrade specimens, *Eukoenenia florenciae* (Rucker) from Val Verde County, Texas. Three universal primer sets were used to target the COI region of arthropods, the 16S rRNA of cyanobacteria, and the ITS region of fungi. Additionally, a blocking primer was designed to prevent amplification of the palpigrade DNA itself. DNA from these specimens was extracted, amplified by PCR, and then sequenced using an Illumina MiSeq platform. Sequences were compared to the NCBI GenBank nucleotide database. The presence of arthropods, cyanobacteria, and fungi supports the premise of generalist feeding habits for this species.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

The micro-whip scorpions or "palpigrades" are assigned to the order Palpigradi in the Class Arachnida (Beccaloni 2009). They were first described by Grassi & Calandruccio (1885) from Italy and assigned to the order Microteliphonida; this name was changed in 1888 by Thorell to Palpigradi (Harvey 2002). This group of minute arthropods was the last arachnid order discovered, remains the least well known group of arachnids, and is thought to be the most primitive (Savory 1964; Kral et al. 2008). There are currently over 90 species comprising six extant genera assigned to two families with a single monotypic fossil genus (Rowland & Sissom 1980; Giribet 2014). They are found in soils or caves from all continents except Antarctica (Harvey 2013). Several authors have described these creatures as "enigmatic" (Harvey 2002; Ferreira et al. 2011; Smrž et al. 2013; Giribet et al. 2014). Due to their minute size, cryptic behavior, and the difficulty of maintaining live specimens in the laboratory, knowledge of these animals' behavior and ecology remains quite limited.

Specimens are very small, ranging from 0.65 mm - 2.4 mm in length. They are lightly pigmented and exhibit a white to pale yellow coloration. Numerous setae that protrude from the body aid this blind arachnid in sensory reception (Savory 1964). They possess a relatively long flagellum that can be observed moving horizontally and vertically in the air as the organism moves about and is thought to aid in stimulus detection (Beccaloni 2009). Another unique feature of this group is revealed by the etymology of its name. The Latin word "*palpo"* translates as "to feel one's way" and "*gradus"* means "step by step". This refers to

the method in which the animal uses the first walking legs to touch in a seemingly sensory fashion and moves along on its pedipalps and other three walking legs (Savory 1964). Palpigrades are the only arachnid group known to exhibit this configuration of limb arrangement in conjunction with its locomotor activities. The chelicerae of the organism are thought to function in prey capture and are used in a comb-like fashion for grooming the legs, pedipalps and flagella. These features contribute to the array of characteristics that make these organisms unique among the minor arachnid groups.

The presence of palpigrade specimens along the upper Devils River in Val Verde County in southwest Texas became known in 2010. The study site is located on the Big Oak River Ranch (BORR) which is owned by Dr. Bob LeGrand, of San Angelo. Collecting efforts were found to be most productive only following periods of moderate to heavy rainfall. Specimens collected from this locality were recently sequenced and included in the first phylogenetic analysis of this order (Giribet et al. 2014). Regions on three genes were examined. These included the complete 18S rRNA gene, 2.2kb of the 28S rRNA gene, and the mitochondrial cytochrome *c* oxidase subunit 1 (COI) gene. Results uindicate the assignment of these Val Verde County specimens to the genus *Eukoenenia* (Grassi & Calandruccio, 1885) and conspecific with three other *Eukoenenia* populations from Mexico; Brazil, and Slovakia. Giribet et al. (2014) support the assignment of these four populations to *Eukoenenia florenciae* (Rucker, 1903) based on near identical sequences for the 18S, 28S, and COI genes (Giribet et al. 2014). This species was first described by Rucker (1903) from specimens collected from soil under a cedar hedge in Bonham, Texas. The presence of these palpigrades in Texas was thought to be the result of human activity (Rucker 1903). This species has apparently spread on other occasions as well and has a worldwide distribution

(Christian & Christophoryova 2013). No males have yet been identified for this species. This complicates an in depth species comparison as it prevents a full characterization of this taxon. Additional collections and further genetic sequencing are necessary to provide a definitive taxonomic assignment of specimens from the Val Verde county population.

Due to the almost universal use of carnivory among other arachnid orders, earlier workers proposed that palpigrades also exhibited predatory feeding habits. This premise is also supported by a report by Conde (1996) in which he noted personal observations by P. Weygoldt of captive palpigrades capturing small collembolans with their chelicerae. A more recent study by Smrž et al. (2013) on *Eukoenenia spelaea* (Peyerimhoff, 1902) from Ardovská Cave in Slovakia has resulted in the need for a further examination of the overall feeding habits of the palpigrades. Seventeen specimens of *E. spelaea* were examined over the course of one year using histological methods. A scanning electron microscope was also used to obtain images for analysis. The results of their study indicate that *E. spelaea* is consuming cyanobacteria. This was determined both by the visual evidence of the single-celled cyanobacteria *Chroococcidiopsis*, as well as evidence of guanine metabolites from cyanobacterial metabolism (Smrž et al. 2013). Results of the study by Smrž et al. (2013) were both unexpected and significant in that it revealed the first evidence supporting a known food source for an individual species of palpigrade.

The specimens from the BORR are found directly in the soil surrounding the river banks near several springs that feed into the upper Devils River. The BORR collection site is a riparian environment partially shaded by large oak trees. This relatively open air habitat is considerably different from the cave habitat of the Smrž et al. (2013). It does not seem likely that the highly adapted cave-dwelling genus of cyanobacteria, *Chroococcidiopsis*, is present

at the BORR collection site but considering that "…cyanobacteria occur virtually in every terrestrial habitat on our planet" (Rindi 2007) there are likely other genera present. Many small arthropods, such as springtails and fungal hyphae have been found with palpigrades at the BORR collection site. If palpigrades feed on any or all of these items, fragments of those items would be present in the palpigrade's digestive tract and could be both detected and identified using molecular techniques.

The advent of more recently developed molecular techniques facilitates an innovative approach to this group's feeding behavior. A recent study by Eitzinger et al. (2013) on centipede dietary analysis demonstrated the use of a PCR assay that targeted 12 different prey species or groups. The gut contents of lithobiid centipedes were analyzed using a PCR assay that was designed to be broad enough to capture the centipede's potentially "generalist" diet but with enough specificity to provide relevant prey information. The primers used in the above study were a combination of species-specific and group-specific assays designed to target prey items and organisms found in the same soil as the centipedes. The primers were developed and then tested on the gut content of the centipede. Results indicated the assay had high specificity and sensitivity to the target species. The present study on *E. florenciae* expands this principle of analysis by Eitzinger et al. (2013) to examine the diet of palpigrades.

The application of PCR assay for understanding the diet of palpigrades is also supported by the study of Remen et al. (2010) on the gut content analysis of oribatid mites. The size of a mite is much closer to the size of a palpigrade and can provide additional insight into the preparation of palpigrade specimens for gut content analysis. Due to the palpigrade's minute size and unknown feeding strategies, this current study would be

expected to encounter obstacles similar to those in the mite study. The mites being tested were known fungivores and therefore had the potential to be contaminated with the targeted food source. To address low DNA amounts and body surface contamination, the authors suggest dissecting out the gut and pooling the content from multiple individuals, and using a wash solution to remove foreign DNA. This current investigation incorporated decontamination and dissection but did not pool gut contents from multiple specimens.

Metabarcoding is a method that allows for rapid detection of species from a mixed sample of DNA sequences (Pompanon et al. 2012; Ji et al. 2013; Clare 2014; Cristescu 2014). This relatively new technique has been utilized for numerous studies, including environmental diversity surveys of plants and animals (Bohmann et al. 2014; Yang et al. 2014; Thomsen & Willerslev 2015; Fahner et al. 2016; Shaw et al. 2016), forensic analysis (Staats et al. 2016), and dietary analysis from fecal pellets (Clare et al. 2009; Hope et al. 2014), gut contents (Leray et al. 2013), and even external sources, such as spider webs (Xu et al. 2015). An increasing number of diet analysis studies have used next generation or high throughput DNA sequencing to analyze the diets of bats (Clare et al. 2009; Hope et al. 2014), invertebrates (Eggs et al. 2013; Saitoh et al. 2016), fish (Leray et al. 2013), as well as microbial populations in humans (Gill et al. 2006) and ruminates (Kartzinel et al. 2016).

The objective of this current study was to test for the presence of DNA from arthropods, cyanobacteria, and fungi within the palpigrade's digestive system using high-throughput sequencing techniques.

MATERIALS & METHODS

Specimen collection.-Collections used in this study were made in June and October of 2015. Sampling included sieving with a wire screen for live specimens. The most productive areas were in the upper several centimeters of the soil, just beneath the leaf litter. Once located, specimens (Table 1) were trapped using a small, damp detailing paintbrush, transferred to a cryotube, and then frozen in liquid nitrogen. Sieving for live specimens proved to be the most effective method of collecting and was consistently used during this study.

Study site.-All collections were made along the banks of a small tributary of the Devils River, near Bakers Crossing in Val Verde County, Texas (29°56′07″N, 101°05′13″W, elevation 1456ft./444m). This area is dominated by karst topography and limestone outcrops. The soil is a “gravelly loamy alluvium” (NRCS Web Soil Survey, https://www.nrcs.usda.gov). Access to the study site (Big Oak River Ranch) was granted by permission of the owner, Dr. Robert LeGrand of San Angelo, Texas.

Preliminary study and results.-A preliminary analysis was conducted to assess the feasibility of the proposed methods. Two palpigrade specimens were decontaminated and DNA was extracted from the entire organism. The DNA concentration was quantified using a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). The DNA samples were then sent to the RTL Genomics Lab, Lubbock, Texas for PCR amplification. Universal primer pairs, ZBJ-ArtF1c/ZBJ-ArtR2c (Zeale et al. 2011), and Cyan359F/Cyan781R (Nübel et al. 1997) were used to target the barcoding regions COI and 16S rRNA respectfully, to test for arthropods and cyanobacteria (Table 2).

Table 1. Samples used to investigate the diet of *Eukoenenia florenciae* in this study and the description of the DNA preparation. All specimens were collected along the bank of a small stream feeding into the Devils River at the Big Oak River Ranch, south of Bakers Crossing, Val Verde County, Texas. Specimen labels reflect the specimen number for that collection followed by the month of the collection in roman numerals and finally the last two digits of the year. For example: P9vi15 was the ninth palpigrade collected in June of 2015. Sample 5 ($P_8W_{R)}$ consisted of the rinse water from sample #4 and was used as a control to test the disinfecting and rinsing methods.

Sample	Label	Date Collected	Description
a	P9vi15	16 June 2015	Disinfected, whole extraction (Preliminary run)
b	P11vi15	16 June 2015	Disinfected, whole extraction (Preliminary run)
1	P3vi15	16 June 2015	Disinfected, whole extraction
2	P4vi15	16 June 2015	Disinfected, whole extraction
3	P7vi15	16 June 2015	Not disinfected, whole extraction
4	P8vi15	16 June 2015	Rinsed with water, whole extraction
5	P_8W_R	(16 June 2015)	Rinse water from #4, Specimen 8
6	P13x15	17 October 2015	Disinfected, dissected, gut extraction
7	P9x15	17 October 2015	Disinfected, dissected, gut extraction
8	P10x15	17 October 2015	Disinfected, dissected, gut extraction
9	P11x15	17 October 2015	Disinfected, dissected, gut extraction
10	P12x15	17 October 2015	Disinfected, dissected, gut extraction

Table 2. Forward and reverse primer sequences used for PCR amplification of gut contents of *Eukoenenia florenciae*. Arthropod-specific primers were used to target the cytochrome oxidase I (COI) gene, ZBJ-ArtF1c/ZBJ-ArtR2c (Zeale et al. 2011) for the preliminary study and LepF1/MLepFF1-Rev for the secondary study (Hebert et al. 2004; Brandon-Mong et al. 2015). The primers Cyan359F and Cyan781R were designed specifically to amplify the 16S ribosomal subunit of cyanobacteria (Nübel et al. 1997). The primers ITS3F and ITS4R, designed by White (1990), were used to amplify the internal transcribed spacer (ITS) operon of fungus. Primers are denoted as forward or reverse by the letters 'F' or 'R' in the name.

Primer	**Sequence (5'- 3')**	**Gene**	**Product (bp)**
ZBJ-ArtF1c	AGATATTGGAACWTTATATTTTATTTTTGG	COI	160
ZBJ-ArtR2c	WACTAATCAATTWCCAAATCCTCC		
LepF1	ATTCAACCAATCATAAAGATATTGG	COI	218
MLepF1-Rev	CGTGGAAAWGCTATATCWGGTG		
Cyan359F	GGGGAATYTTCCGCAATGGG	16S rRNA	379
Cyan781R	GACTACWGGGGTATCTAATCCCWTT		
ITS3F	GCATCGATGAAGAACGCAGC	ITS rRNA operon	330
ITS4R	TCCTCCGCTTATTGATATGC		

*(3 = with a Spacer C3 at the 3'end)

The samples were sequenced using the Illumina MiSeq platform and the sequences were processed through the RTLGenomics bioinformatics pipeline (Research and Testing Laboratories, Data Analysis Methodology version 2.2.4). The results warranted the need for a blocking primer and inclusion of the internal transcribed spacer (ITS) region of the fungus genome as an additional target gene. An annealing blocking primer is designed to be species specific and prevents the study species from amplifying (Vestheim & Jarmon 2008; Leray et. al 2013) allowing other arthropods that might be present in the gut to amplify instead.

Specimen decontamination.-Prior to DNA extraction, specimens were decontaminated. The specimens were washed in a 2.5% bleach solution (Greenstone et al. 2005) for 15 minutes on an auto-rocker. They were then rinsed for 15 minutes using distilled water to prevent bleach from inhibiting further reactions. Two samples were not washed in order to test the effect of the decontamination process on subsequent DNA sequence results. One of these samples was instead rinsed three times with distilled water. This rinse water was then treated as a sample and run through the extraction, PCR, and sequencing steps. The other sample was not rinsed in bleach or water, allowing for any foreign DNA on the external surface of the palpigrade to be amplified.

Specimen dissection.-Due to their very small size, initial attempts to dissect and remove the digestive tract from the palpigrade were unsuccessful. Instead, the entire palpigrade specimen was used for each DNA extraction. Successful dissections later allowed for five samples of isolated digestive tract that were extracted, amplified, and then sequenced. The dissection protocol established is summarized below. All tools were sterilized in a UV CrossLinker (Thermo Fisher Scientific, Waltham, MA) for three minutes on each side and surfaces were disinfected with 10% bleach solution before use. Following

disinfection the specimens were transferred to sterile Phosphate Buffered Saline (PBS) solution (Andrews 2013). A microdissection dish was made with black silicone sealant in a small plastic petri dish and dissections were completed under a Wild Heerbrugg Stereo dissecting microscope (Wild Heerbrugg AG, Heerbrugg, Switzerland) with the aid of external light sources. Three minuten insect pins (BioQuip, Rancho Dominguez, CA, USA) were inserted into the specimen, at the prosoma and both ends of the opisthosoma to hold it in place. The cuticle of the opisthosoma was then split down the mid-sagittal plane with a tungsten needle and pinned back with two minuten pins (BioQuip, Rancho Dominguez, CA, USA). The contents within the opened opisthosoma were scraped out and pipetted into a microcentrifuge tube for DNA extraction. The cuticle and remaining structures of the palpigrade were transferred to 70% ethanol for permanent storage.

Blocking primer design.-For whole specimen DNA extraction, the gut contents comprise a very small percentage of the genetic material present in the DNA sample. Consequently, it was possible that any DNA from food items might not be amplified because of the overwhelming presence of the palpigrade DNA. In the preliminary run, 99.68% of the sequence reads were those of *E. florenciae*. In order to decrease the amount of palpigrade DNA present and increase the likelihood that any prey/food DNA would amplify, a species-specific annealing blocking primer was designed to block the palpigrade DNA. An annealing blocking primer binds to the target region, overlapping the reverse primer binding site and blocking it from annealing. Since the reverse primer can no longer bind to the DNA, that fragment will not be amplified and the amount of palpigrade DNA sequence that is produced is reduced (Vestheim & Jarmon 2008; Leray et al. 2013).

The COI universal barcode primers ZBJ-ArtF1c/ZBJ-ArtR2c (Zeale et al. 2011) were aligned to a 654 bp COI sequence for *E. florenciae* (GenBank entry KF 823880.1) in MEGA7 (Kumar et al. 2016) using ClustalW and adjusted by eye. This sequence was selected because it was from a specimen collected at the same locality as this current study. Two barcode sequences from the preliminary run were then also aligned to the GenBank reference sequence of *Eukoenenia florenciae* to confirm that the primer position was correct. The reverse complement of the reverse primer (MLepF1-Rev) was then aligned to the GenBank reference sequence to identify an appropriate location for an annealing blocking primer. The blocking primer was designed following the criteria outlined in Leray et al. (2013). Several COI sequences from arthropods (n=66) were downloaded from GenBank. The sequences chosen were the top 45 GenBank matches to *Eukoenenia florenciae* as well as species commonly seen in the litter samples from the study site. The sequences were aligned in MEGA7 (Sudhir et al. 2016) using ClustalW and adjusted by eye.

The entropy (Hx) was calculated using the BioEdit software program, version 7.2.5 (Hall 1999). This program evaluated the nucleotide composition at each base pair site. Sites that are conserved (the base is the same for all sequences) have an entropy value of 0, sites at which all nucleotides are equally represented (with a frequency of 0.25) have an entropy value of 1. This process identified sites that were highly dissimilar within the universal primer binding site in order to decrease the likelihood that the blocking primers would bind to non-target DNA and would preferentially bind to palpigrade DNA. The plots were evaluated for a site of maximal entropy at the 3’ end approximately 15-19 bp “upstream” from the start of the universal primer (Figure 1). Blocking primers were designed at potential sites as the reverse complement of the reference sequence. These were then evaluated in

OligoAnalyzer (available free on www.idtdna.com) for GC content, melting temperature, and secondary structures. The blocking primer chosen, Eflorenciae_Block, 5'ATGCTATATCAG GACATCCTAATATTAAG3* 3' was the optimal combination of all criteria considered (the character, 3*, represents a C3 spacer added to the 3' end). It starts at site 204, has an overall length of 29 bp and overlaps the universal primer by 15bp (Figure 1).

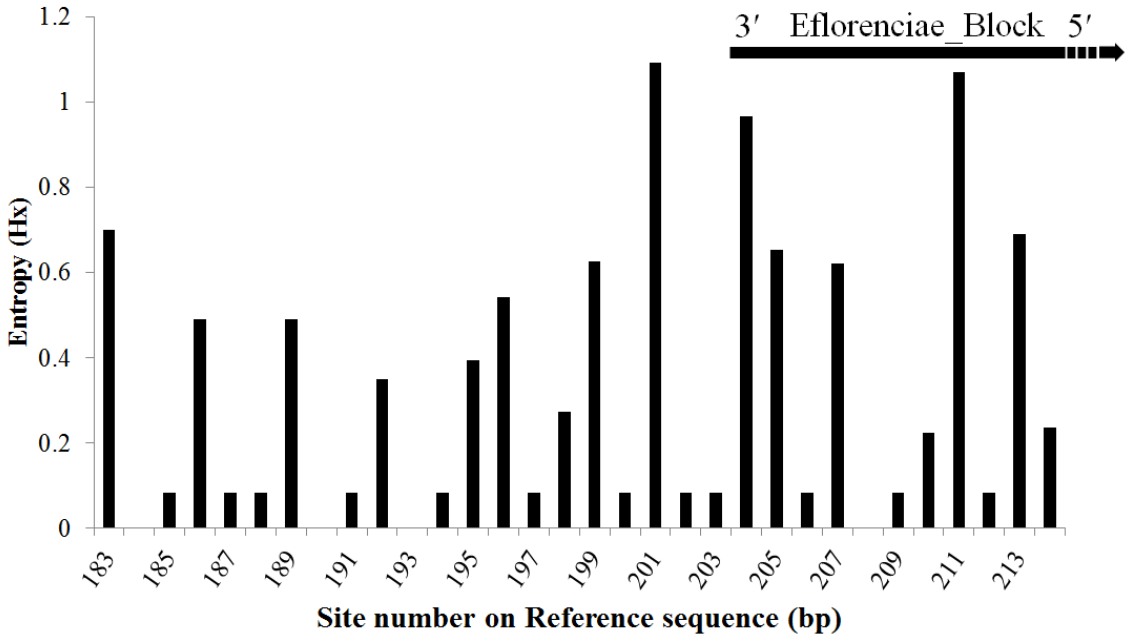

Figure 1. Entropy at each site for an alignment of 66 arthropod COI barcode region sequences. The sequences chosen for the alignment were highly similar to or known to occur with the Val Verde County population of *Eukoenenia florenciae*. Graph shows a 71 bp region downstream of the universal primer, MLepF1-Rev, binding site. The blocking primer chosen, Eflorenciae_Block, starts at site 204 and spans 29 base pairs, overlapping the 3′ end of the MLepF1-Rev primer (starting at site 218) by 15bp.

DNA extraction and PCR amplification.-Specimens were extracted following the Blood and Tissue protocol (07/2006) of the Qiagen DNeasy extraction kit (Valencia, CA) resulting in two elutions. The DNA concentration was measured using a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA). The elution of extracted DNA with the highest concentration was sent to the Genomics Core lab at Texas A&M Corpus Christi. The COI primers, LepF1/MLepF1-Rev (Brandon-Mong et al. 2015) were chosen instead of ZBJ-ArtF1c/ZBJ-ArtR2c for the final portion of this study. This change was supported because these new primers were found to be more effective at detecting arthropods when compared to several other COI barcoding primers and their target region encompassed the 157bp that was amplified during the preliminary part of this study (Brandon-Mong et al. 2015). The 16S rRNA region was amplified using the primer pair Cyan359F/Cyan 781R (Nübel et al. 1997) and the internal transcribed spacer (ITS) was amplified using the primer pair ITS3F/ITS4R (White 1990). The conditions and reagent concentrations used for PCR amplification are summarized in Appendix 1. Following amplification the reactions were combined for sequencing.

DNA sequencing.-The samples were sequenced on an Illumina Miseq. This platform uses reversible-terminator sequencing by synthesis technique (McElhoe et al. 2014). Paired end sequencing (2x250) was used which sequences the fragments in both the 3' to 5' direction and the 5' to 3' direction, allowing for a higher quality read (Shendure 2008; Ansorge 2009; McElhoe et al. 2014). Both the preliminary and final datasets were sequenced using the Illumina Miseq platform with slight differences in how the samples were organized. The final dataset was processed through the Bioinformatics pipeline at the Texas A&M Corpus Christi Genomics Core Lab (C. Bird, Director, Genomics Core Lab, Corpus Christi, Texas).

Data analysis and taxonomic assignment.-Sequences were first separated by each genetic marker (ITS, COI, 16S). Sequences were then filtered to remove PCR errors, low quality sequence reads, and chimera sequences. Operational taxonomic units (OTUs) were created by clustering sequences with 97% similarity (Hao et al. 2011). A representative sequence from each OTU was then queried on GenBank for taxonomic assignment. Results were filtered down to those that had a query cover of at least 80% and a species identity of at least 97%. For a more stringent evaluation, a filter of 100% query cover, 97% identity and 100% query cover, 98% identity were also applied. The initial query cover criteria of 80% was chosen based on recommendation from the Core Genomics Lab. A 97% identity match is a generally accepted value for species level identification (Leray & Knowlton 2015).

RESULTS

Arthropods.-The presence of *Eukoenenia florenciae* and other arthropod sequences were confirmed by results from the COI gene. From the two preliminary samples there were 53,866 quality sequence reads from the COI dataset. In this dataset, there were 28 total OTUs, three of them were less than 160bp and did not match any entries in GenBank, one OTU with 4 sequence reads in sample a (P9vi15) had a low match to a fungus species, one OTU with 18 sequences from sample a and 15 sequences in sample b (P11vi15) matched equally to three aphid species and the remaining 23 OTUs, with 53,805 sequences between both samples, matched to *E. florenciae*.

In the COI dataset from the secondary study, there were a total of 8,843 sequence reads divided into 1,162 OTU clusters. Filtering down to those with a query cover ≥ 80% and a BLAST sequence ID ≥ 97% narrowed the dataset to 4,524 reads and 89 OTUs. These 89 OTUs matched ten unique species: *Eukoenenia florenciae*, *Blatella germanica* (German cockroach), *Harmonia axyridis* (Harlequin ladybird beetle), *Zelus renardii* (Leafhopper assassin bug), *Elephas maximus* (Asian elephant), *Cerebratulus longiceps* (ribbon-worm), *Coccinella spetempunctata* (Seven-spot Ladybird beetle), *Olcella trigramma* (grass fly), *Acheta domesticus* (House cricket), and *Anser canagica* (Emperor goose). OTUs matching to *Eukoenenia florenciae* were present in all but one sample. *Anser canagica* and *Elephas maximus* are likely contaminants and *Cerebratulus longiceps* is likely a misidentification on GenBank. These three artifacts are further explicated within the discussion. The other species are all small arthropods with distributions encompassing the study site. Ecologically, they could feasibly be considered food items. Three species: *Blatella germanica, Zelus renardii* and *Coccinella spetempunctata* were each only found in one sample. *Acheta domesticus* was

found in two samples and three species, *Harmonia axyridis, Olcella trigramma,* and *Cerebratulus longiceps,* were each repeated in three samples.

When the filtering criteria was increased to 100% query cover, 97% species identity, five of the insect species: *Blatella germanica, Coccinella spetempunctata, Harmonia axyridi, Acheta domesticus, Olcella trigramma,* as well as *Anser canagica, Cerebratulus longiceps,* and *Eukoenenia florenciae* remained. This more stringent criterion adds further confidence to the taxonomic assignments.

Cyanobacteria.-Bacterial DNA was detected using the 16S primer assay. In the preliminary 16S dataset, there were 114 quality sequence reads. These divided into two OTUS. One OTU had six sequences (from sample b) that matched equally to several chloroplast entries. The other OTU had 108 sequences (from sample b) that matched equally to a cyanobacteria, *Arthrospira platensis,* an uncultured bacterium clone, and a chloroplast from the green algae genus *Scenedesmus*.

From the secondary study, there were a total of 613 sequence reads divided into 346 OTU clusters. Filtering down to those with a query cover ≥ 80% and a BLAST sequence ID ≥ 97% narrowed the dataset to 31 reads and 18 OTUs. The majority of these reads were found in sample 3, the unwashed specimen. The bacteria species matched to both Gram positive and Gram negative bacteria: *Pseudonocardia endophytica, Pseudonocardia petroleophila, Bacillus cereus, Bacillus cihuensis,* and *Staphylococcus capitis* which are commonly found in the environment (Table 3). There were also sequence reads found in two other samples. However, a confident taxonomic identification could not be made for these due to the fact that they matched to several database entries to eukaryotic chloroplast and cyanobacteria sequences equally.

Fungus.-The fungal ITS region was not targeted in the preliminary study, therefore all ITS results are only from the final study. There were a total of 3,396 sequence reads divided into 2,380 OTU clusters. Filtering down to those with a query cover ≥ 80% and a BLAST sequence ID ≥ 97% narrowed the dataset to 24 reads and 11 OTUs. The sequences remaining after filtering were from only three samples; remaining samples did not have any sequences with matches to a species entry on GenBank with greater than 97% identity. From the OTUs, four species of fungus were identified: *Talaromyces amestolkiae, Lecanicillium* sp. (equivalent match to both *L. saksenae* and *L. psalliotate*), *Cladosporidium sphaerospermum* and *Malasseiza restricta.* Each of these species occurred in two samples and *Cladosporidium sphaerospermum* occurred in three samples.

Talaromyces amestolkiae is a common cosmopolitan sac fungus that has been cultured from both indoor and outdoor locations. *Talaromyces amestolkiae* and other members of the *Talaromyces* genus are common in soil (Yilmaz et al. 2014). Members of the *Lecanicillium* genus are entomopathic fungi. They naturally infect insects and are commonly used in commercial pesticides. *Cladosporidium sphaerospermum* is a common mold species that can be found on living and dead plant material (Zalar et al. 2007). OTUs that originally matched to “Fungal endophyte” entries also had a high, 99%, identity match to this species. *Malassezia restricta* is a Basidiomycota that is found on the human skin.

Table 3. Taxonomic identification of sequences from molecular diet analysis from specimens of the palpigrade *Eukoenenia florenciae.* Samples a, b, 1, and 2 are disinfected, whole palpigrades; sample 3 was not disinfected, sample 4 was rinsed with sterile water, sample 5 was the collected rinse water used on sample 4, and samples 6-10 were the dissected digestive tracts from disinfected palpigrades. Species identification of the OTUs (Operational Taxonomic Units) that had query coverage greater than or equal to 80% and identity match greater than or equal to 97%.

Sample	COI	16S	ITS
a	*Eukoenenia florenciae*	*Arthrospira platensis*	-
b	*Eukoenenia florenciae*	-	-
	Aphis spp. (*Aphis craccivora*, *Aphis spiraecola*, *Aphis rumicus*)	-	-
1	*Eukoenenia florenciae*	-	-
	Blatella germanica	-	-
	Harmonia axyridis	-	-
	Zelus renardii	-	-
	Elephas maximus	-	-
	Cerebratulus longiceps	-	-
2	*Eukoenenia florenciae*	Unidentified, entry FQ683742.1	-
	Coccinella septempunctata		
	Olcella trigramma		
3	*Eukoenenia florenciae*	*Pseudonocardia endophytica*	-
	Olcella trigramma	*Pseudonocardia petroleophila*	
	Cerebratulus longiceps	*Bacillus cereus*	
		Bacillus cihuensis	
		Staphylococcus capitis	
		Granulicatella para-adiacens - JF803551.1	
		bacterium WX65 - KC921189.1	
		Unidentified - FQ683641.1	
		unidentified bacterium - EF220533.1	
4	*Eukoenenia florenciae*	-	*Talaromyces amestolkiae*
			Lecanicillium saksenae
			Malassezia restricta
			Fungal endophyte - KR017021.1
5	*Eukoenenia florenciae*	-	-
	Anser canagica		
6	*Eukoenenia florenciae*	-	*Cladosporium sphaerospermum*
	Cerebratulus longiceps		*Malassezia restricta*
			Fungal endophyte- KR017021.1
7	*Harmonia axyridis*	-	-
	Acheta domesticus		
8	*Eukoenenia florenciae*	*Virgulinella fragilis*	-
		bacterium EA10-B11-13 - JF418020.1	
9	*Eukoenenia florenciae*	-	-
10	*Eukoenenia florenciae*	-	*Cladosporium sphaerospermum* *Talaromyces amestolkiae*
	Harmonia axyridis		
	Olcella trigramma		
	Acheta domesticus		

DISCUSSION

The primary objective of this current study was to test for the presence of DNA from arthropods, cyanobacteria, and fungi within the palpigrade's digestive system using high-throughput sequencing techniques. Arachnids are known to predigest food items externally (Beccaloni 2009). Therefore, the DNA of food items that are consumed are already degraded before being ingested (Hereward & Walter 2012) but they are still expected to amplify for the relatively short (< 350 bp) barcoding regions that are targeted (Symondson 2002; Jurado-Rivera et al. 2009). This current palpigrade study found that using both whole-extracted and gut-dissected specimens, miniscule amounts of partially digested food items could produce a DNA barcode that could be matched to a database. Sequences matching with an 80% query cover and 97% identity to organisms from all three proposed groups were identified. Although some studies exclude rare prey items from analysis (Brown et al. 2014), they can still provide information about the diet and its environmental implications (Clare 2014). Because there is very little known about the feeding habits of this group, all information is potentially valuable and this current study reports all taxa. More samples should be analyzed to increase the confidence in the actual presence of the sequences and a better understanding of the full breadth of the palpigrade's diet.

The species specific blocking primer, Eflorenciae_Block, was found to be effective in the secondary study, reducing the percentage of palpigrade sequences from 99.76% in the primary run to 42.22% in the final run. Because other conditions (different COI primers, the addition of the ITS primers, dissection of some specimens, and a different sequencing facility), were altered between primary and secondary runs, this decrease cannot confidently

be attributed solely to the blocking primer but it seems likely to have had a substantial contribution.

Feeding method.-The adult stages of all the arthropod species identified are much larger than the palpigrades. In addition to the size difference, the thick protective elytra of the ladybird beetles, and the mobility of the fly, cricket and assassin bug would likely prevent the miniscule and more fragile palpigrade from preying on these organisms. Therefore if the palpigrade is in fact feeding on the adult forms of these species it is likely scavenging fragments of dead organisms in the soil. However, the larval stages or eggs of these insects have a much thinner exoskeleton or cuticle than in adult forms and are closer in size to palpigrades. They could feasibly be captured by the palpigrade. More observational studies appear warranted to address just how palpigrades are consuming the arthropods.

Limitations of the study.-Much of the difficulty in the collection, dissection, and amplification of specimens is due to their very small size and elusive nature. Specimens which were returned to the laboratory for observation were unable to survive the artificial laboratory conditions. It appears that the optimal conditions required for survival are within very narrow limits. Based on observations of the natural environment, these abiotic conditions likely include a relatively high humidity with warm temperatures and low light levels. Observations of additional feeding behaviors are dependent upon additional efforts in maintaining living specimens in a laboratory setting. If possible, controlled laboratory feeding studies followed by sequencing would provide justification of the results presented here.

Although this investigative technique is potentially very revealing, its limitations should also be addressed. A common recurring issue in molecular diet analysis studies is the

difficulty in distinguishing between primary prey and secondary prey species (Harwood et al. 2001; Sheppard et al. 2005). For example, the assassin bug, *Zelus renardii*, is a known predator that feeds on other small insect eggs, larvae, and adults (Drees & Jackman 1998). Therefore the origin of a DNA sequence remains in question; it may represent a primary prey item or may have first been ingested another organism that was then ingested by the palpigrade.

The state of the palpigrade at the time of collection must also be considered. Results could vary depending on how much time has passed since the palpigrade's last feeding activity. The condition of the prey/food DNA degrades as digestion progresses and could limit the detection. This is especially relevant when considering that the digestion process begins externally.

Database bias.-At this time, the NCBI GenBank houses nucleotide sequence entries for over 300,000 species. However, there are still many organisms that are not represented or identified, especially at lower taxonomic levels (Leray & Knowlton 2016). This seems to be particularly true for minute, soil-dwelling arthropods as well as bacteria and fungi. This issue can result in some degree of database bias during taxonomic assignment. If there is no entry in GenBank for the queried species then one cannot determine a species match. Any higher level taxonomic identification would depend on the relative similarities to sequences that are available in the database.

Additionally, for the 16S and ITS databases, there are large amounts of "unknown" or "environmental sample" database entries. When a queried sequence matches these "unknown" entries, a species-level identification is not possible. If there is published documentation to support those sequences, then this information could still provide valuable

insight from their locality data or higher level taxonomic identification. The challenge, however, is that it remains possible to publish a sequence on GenBank without referencing the study from which it originated. Therefore many of these "unknown" or "environmental" sequences are from unpublished studies and have no accessible supporting information.

High throughput sequencing has been shown to be able to sequence DNA in very small relative concentrations, "from a few nanograms to tens of micrograms" (Loman et al. 2012), as well as DNA that is degraded or archived for a long period of time (Hykin et al. 2015). Due to this sensitivity, it is also necessary to acknowledge the possibility of environmental or technician contaminants. Results of this study yielded matches to elephant, ribbon worm, and goose DNA sequences, which were unexpected and not found at the study site. It is highly unlikely that these items represent prey items of the palpigrades. The elephant DNA was almost certainly a contamination. After initial confusion it was realized that the DNA was likely introduced after handling stored museum specimens, including parts from an Asian elephant that originated from the San Antonio Zoo. This occurred the same day as the DNA extraction of the sample and apparently DNA was transferred from the author's contaminated garments or gloves to the sample tube. The sequences were only recovered from one sample and were not present in the dataset filtered with the highest stringency (100% query cover, 97% sequence identity). The DNA that matched to goose was found only in the rinse water from the specimen that was not treated with a bleach solution. Upon further investigation it was revealed that the sequences actually matched equally to the common domestic chicken, *Gallus gallus.* There is not enough resolution to determine which fowl species it is but since these are both fairly common, and it only occurred in one of the unwashed samples, it is clearly a contaminant.

The other unexpected identification was to a species of marine ribbon worm, *Cerebratulus longiceps.* Because the study site is not in or around a marine habitat this was not a feasible food item or a likely contaminant. It was also found in three samples, indicating that it was an actual diet item. The first explanation considered was that it could actually belong to a terrestrial ribbon worm. Although they are not known to exist in Val Verde County, it could be possible that there was a new species there. To test this idea, the sequence was again queried using BLAST on GenBank, within the available sequences for *Geonemertes*, the only genus of terrestrial ribbon worms. However, there were no strong matches. The next approach was to test the validity of the identification of the GenBank sequences. When the diet sequence was queried using BLAST on GenBank, it matched with 100% query cover and high identity 97-99% to four *Cerebratulus* environmental sample clones (JQ007428-JQ007431) and *Cerebratulus longiceps* (EF124987.1). The four environmental samples (JQ007428-JQ007431) were derived from dust particles of an ancient shroud (Barcaccia et al. 2015). The authors found that the sequences were "ascribable" to the marine ribbon worm *Cerebratulus longiceps* based on a strong match to the GenBank entry EF124987.1. According to the NCBI GenBank website (www.ncbi.nlm.nih.gov), the entry EF124987.1 is from an unpublished study entitled Molecular Phylogeny of Pilidiophoran Nemerteans. Therefore there was no additional information regarding the basis for the identification of sequence EF124987.1. Next, the sequence (EF124987.1) was queried using BLAST in GenBank. If it was a ribbon worm it would be expected that there would be at least low matches to other ribbon worm species (There are no other GenBank entries for the COI gene of *Cerebratulus longiceps* but there are at least 88 COI entries for the genus *Cerebratulus*). Top results from the query did not match any ribbon worm species. Instead

the top four matches (below the environmental samples JQ007428-JQ007431) followed by the percent query cover and identification were: *Fujientomon dicestum* (proturan), HQ882827.1, 99%|88%, Cyclopoida environmental sample clones (copepod), KP136566.1, 99%|88%; KP136575.1 99%|88%, and *Culex bahamensis* (mosquito) JX260644.1, 63%|99%. All other results generated were below 50% query cover. These results suggest that the *Cerebratulus longiceps* entry EF124987.1 is not actually a ribbon worm but is another arthropod. The taxonomy of these sequences could be determined by collecting, morphologically identifying, and then sequencing other arthropods from the study site. Following identification, the original authors of the GenBank sequence should be contacted in order to change the identity of the existing GenBank entries. GenBank sequence errors in quality and sequence identity are a widespread problem (Bridge et al. 2003; Forster 2003; Harris 2003) and have been the focus of multiple studies (Longo et al. 2011; Spouge & Mariño-Rameríz 2012; Shen et al. 2013) in order to make researchers aware of this problem as well as to prevent and identify these errors.

Contamination was also a concern due to the small amounts of DNA being tested. Work areas and all utensils were treated with care and sterilized before use. However, it has been reported that reagents from prepackaged kits, such as extraction buffers, etc. have been the source of contamination for high-throughput sequencing studies (Glassing et al. 2016). This has been reported primarily as bacterial contamination. This would potentially cause concern for the 16S dataset but did not seem to be an issue in this study. The types of bacteria typically responsible for this contamination were only observed in a single sample. This sample contained the palpigrade specimen that was not washed or rinsed prior to DNA

extraction. This sample had the highest number of OTUs from the 16S dataset and most likely originated from the natural bacterial community on the surface of the palpigrade.

This study serves as a basis for future study on this topic. Sequences from cyanobacteria, arthropods and fungi were detected from within the palpigrade. Additional samples for molecular analysis and further field studies should be completed for a more thorough understanding of the extent of diversity in the diet and the actual feeding methods. Additional work is needed to determine if the arthropods that are being consumed are in the adult or larval form. Culturing samples of fungi and cyanobacteria from various locations in the habitat could be sequenced, identified, and then compared to the results of the diet analysis. Having a database of sequences from the study site would assist greatly in the assignment of prey items. If there is a difference in where the species that were consumed are found, i.e. throughout the soil or on decomposing arthropods, this could indicate whether they are being consumed intentionally or incidentally.

LITERATURE CITED

Andrews, E. 2013. Analyzing arthropods for the presence of bacteria. Current Protocols in Microbiology. 28:1E.6.1-1E.6.14.

Ansorge, W. 2009. Next-generation DNA sequencing techniques. New Biotechnology 25:195-203.

Barcaccia, G., G. Galla, A. Achilli, A. Olivieri, & A. Torroni. 2015. Uncovering the sources of DNA found on the Turin Shroud. Scientific Reports 5:14484.

Beccaloni, J. 2009. Arachnids. University of California Press. Berkeley and Los Angeles, California.

Bohmann, K., A. Evans, M.T.P. Gilbert, G.R. Carvalho, S. Creer, M. Knapp, D.W. Yu, & M. de Bruyn. 2014. Environmental DNA for wildlife biology and biodiversity monitoring. Trends in Ecology & Evolution 29:358-367.

Brandon-Mong, G.J., H.M. Gan, K.W. Sing, P.S. Lee, P.E. Lim, & J.J. Wilson. 2015. DNA metabarcoding of insects and allies: and evaluation of primers and pipelines. Bulletin of Entomological Research 105:717-727.

Bridge, P.D., P.J. Roberts, B.M. Spooner, & G. Panchal. 2003. On the unreliability of published DNA sequences. New Phytologist 160:43-48.

Brown, D.S., W.O.C. Symondson, R. Burger, N. Cole, D. Vencatasamy, E.L. Clare, & A. Montaxam. 2014. Dietary competition between the alien Asian Musk Shrew (*Suncus murinus*) and a reintroduced population of Telfair's Skink (*Leiolopisma telfairii*). Molecular Ecology 23:3695–3705.

Clare, E.L. 2014. Molecular detection of trophic interactions: emerging trends, distinct advantages, significant considerations and conservation applications. Evolutionary Applications 7:1144-1157.

Clare, E., E.E. Fraser, H.E. Braid, M.B. Fenton, & P.D.N. Hebert. 2009. Species on the menu of a generalist predator, the eastern red bat (*Lasiurus borealis*): using a molecular approach to detect arthropod prey. Molecular Ecology 18:2532-25422.

Conde, B. 1996. Les Palpigrades, 1885-1995. Acquisitions et lacunes. revue suisse de zoologis, hors-serie 1:87-106.

Christian, E., & J. Christophoryova. 2013. *Eukoenenia florenciae* (Arachnida: Palpigradi): Lessons from a newcomer to Central Europe and the island of Tenerife. Biologia 68:1182-1188.

Cristescu, M.E. 2014. From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. Trends in Ecology & Evolution 29:566-571.

Drees, B.M., & J.A. Jackman. 1998. A field guide to common Texas insects. Gulf Publishing Company. Houston, Texas.

Eggs, B., & D. Sanders. 2013. Herbivory in spiders: The importance of pollen for orb-weavers. PLoS ONE 8:e82637.

Eitzinger, B., A. Micic, M. Korner, M. Traugott, & S. Scheu. 2013. Unveiling soil food web links: New PCR assays for detection of prey DNA in the gut of soil arthropod predators. Soil Biology and Biochemistry 57:943-945.

Fahner, N.A. S. Shokralla, D.J. Baird, & M. Hajibabaei. 2016. Large-scale monitoring of plants through environmental DNA metabarcoding of soil: recovery, resolution, and annotation of four DNA markers. PLoS ONE 11:e0157505.

Ferreira, R.L, M.F.V.R. Souza, E.O. Machado, & A.D. Brescovit. 2011. Description of a new *Eukoenenia* (Palpigradi: Eukoeneniidae) and *Metagonia* (Aranae: Pholcidae) from Brazilian caves, with notes on their ecological interactions. Journal of Arachnology 39:409-419.

Forster, P. 2003. To err is human. Annals of Human Genetics. 67:2-4.

Gill, S.R., M. Pop, R.T. DeBoy, P.B. Eckburg, P.J. Turnbaugh, B.S. Samuel, J. I. Gordon, D.A. Relman, C.M. Fraser-Liggett, & K.E. Nelson. 2006. Metagenomic analysis of the human distal gut microbiome. Science 312:1355-1359.

Giribet, G., E. McIntyre, E. Christian, L. Espinasa, R.L. Ferreira, O.F. Francke, M.S. Harvey, M. Isaia, L. Kovác, L. McCutchen, M.F.V.R. Souza & M. Zagmajster. 2014. The first phylogenetic analysis of Palpigradi (Arachnida) – the most enigmatic arthropod order. Invertebrate Systematics 28:350–360.

Glassing, A., S.E. Dowd, S. Galandiuk, B. Davis, & R.J. Chiodini. 2016. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. Gut Pathogens 8:24.1-24.12.

Grassi, B., & S. Calandruccio. 1885. Intorno ad un nuovo Aracnide Artrogastro (Koenenia mirabils [sic]) che crediamo rappresentante d.un nuovo ordine (Microteliphonida). Naturalista Siciliano 4: 127.133, 162.168.

Greenstone, M., D. Weber, T. Coudron, M. Payton, & J. Hu. 2005. Removing external DNA contamination from arthropod predators destined for molecular gut-content analysis. Molecular Ecology Resources 12:464-469.

Hall, T. 1999. BioEdit: a user friendly biological sequence alignment editor & analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41:95-98. Version 7.2.5. Updated 12/13/2013. http://www.mbio.ncsu.edu/bioedit/bioedit.htmL.

Hao, X., R. Jiang, & T. Chen. 2011, Clustering 16 RRNA for OTU prediction: a method of unsupervised Bayesian clustering. Bioinformatics 27:611-618.

Harris, D.J. 2003. Can you bank on GenBank? Trends in Ecology and Evolution. 18:317-319.

Harvey, M.S. 2013. Palpigrades of the World, version 1.0. Western Australian Museum, Perth. http://www.museum.wa.gov.au/catalogues/palpigrades

Harvey, M.S. 2002. The neglected cousins: What do we know about the smaller arachnid orders?. Journal of Arachnology 30:357-372.

Harwood, J.D., S.W. Phillips, K.D. Sunderland, & W.O.C. Symondson. 2001. Secondary predation: quantification of food chain errors in an aphid-spider-carabid system using monoclonal antibodies. Molecular Ecology 10:2049-2057.

Hebert, P.D.N., A. Cywinska, S.L. Ball, & J.R. deWaard. 2003. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London 270:313-321.

Hereward, J.P., & G.H. Walter. 2012. Molecular interrogation of the feeding behavior of field captured individual insects for interpretation of multiple host plant use. PLoS ONE 7:e44435.

Hope, P.R., K. Bothmann, M.T.P. Gilbert, M.L. Zepeda-Mendoza, O. Razgour, & G. Jones. 2014. Second generation sequencing and morphological faecal analysis reveal unexpected foraging behavior by *Myotis nattereri* (Chiroptera, Vespertilionidae) in winter. Frontiers in Zoology 11:39.1-39.15.

Hykin S.M., K. Bi, & J.A. McGuire. 2015. Fixing formalin: A method to recover genomic-scale DNA sequence data from formalin-fixed museum specimens using high-throughput sequencing. PLoS ONE 10:e0141579.

Ji, Y., L. Ashton, S.M. Pedley, D.P. Edwards, Y. Tang, A. Nakamura, R. Kitching. P.M. Dolman, P. Woodcock. F.A. Edwards, T.H. Larsen, W.W. Hsu, S. Benedick, K.C. Hamer, D.S. Wilcove, C. Bruce, X. Wang, T. Levi, M. Lott, B.C. Emerson, & D.W. Yu. 2013. Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. Ecology Letters 16:1245-1257.

Jurado-Rivera, J.A., A.P. Vogler, C.A.M. Reid, E. Petitpierre, & J. Goméz-Zurita. 2009. DNA barcoding insect-host plant associations. Proceedings of the Royal Society B 276:639-648.

Kartzinel, T.R., P.A. Chen, T.C. Coverdale, D.L. Erickson, W.J. Kress, M.L. Kuzmina, D.I. Rubenstein, W. Wang, & R.M. Pringle. 2015. DNA metabarcoding illuminates dietary niche partitioning by African large herbivores. Proceedings of the National Academy of Sciences. 112:8019-8024.

Kumar, S., G. Stecher, & K. Tamura. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33:1870-1874.

Kral, J., L. Kovac, F. Stahlavsky, P. Lonsky & P. Luptacik. 2008. The first karyotype study in palpigrades, a primitive order of arachnids (Arachnida: Palpigradi). Genetica 134:79-87.

Leray, M., N. Agudelo, S.C. Mills, & C.P. Meyer. 2013. Effectiveness of annealing blocking primers versus restriction enzymes for characterization of generalist diets: unexpected prey revealed in the gut contents of two coral reef fish species. PLoS ONE 8:e58076.

Leray, M., & N. Knowlton. 2015. DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. Proceedings of the National Academy of Sciences 7:2076-2081.

Leray, M., & N. Knowlton. 2016. Censusing marine eukaryotic diversity in the twenty-first century. Philosophical Transactions of Royal Society B 371:20150331.

Loman, N.J., C. Constantinidou, J.Z.M. Chan, M. Halachev, M. Sergeant. C.W. Penn, E.R. Robinson, & M.J. Pallen. High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. Nature Reviews, Microbiology 10:599-606.

Longo, M.S., M.J. O'Neill, R.J. O'Neill. 2011. Abundant human DNA contamination identified in non-primate genome databases. PLoS ONE 6:e16410.

McElhoe, J., M. Holland, K. Makova, M. Shu-Wei Su, I. Paul, C. Baker, S. Faith, & B. Young. 2014. Development and assessment of an optimized next-generation DNA sequencing approach for the mtgenome using the Illumina MiSeq. Forensic Science International: Genetics 13:20-29.

Nübel, U., F. Garcia-Pichel, & G. Muyzer. 1997. PCR primers to amplify 16S rRNA genes from cyanobacteria. Applied Environmental Microbiology 63:3327-3332.

Pompanon, F., B.E. Deagle, W.O.C. Symondson, D.S. Brown, S.N. Jarman, & P. Taberlet. 2012. Who is eating what: diet assessment using next generation sequencing. Molecular Ecology 21:1931-1950.

Remen, C., M. Kruger, & A. Cassel-Lundhagen. 2010. Successful analysis of gut contents in fungal-feeding oribatid mites by combining body-surface washing and PCR. Soil Biology and Biochemistry 42:1952-1957.

Rindi, F. 2007. Diversity, distribution and ecology of green algae and cyanobacteria in urban habitats. Chapter: Algae and cyanobacteria in extreme environments. Volume 11 of the series Cellular origin, life in extreme habitats and astrobiology. pp. 619-638.

Rowland, J.M., & W.D. Sissom. 1980. Report on a fossil palpigrade from the Tertiary of Arizona, and a review of the morphology and systematics of the order (Arachnida). Journal of Arachnology 8:69–86.

Rucker, A. 1903. A new Koenenia from Texas. Quarterly Journal of Miscrocopical Science 47:215-231.

Saitoh, S., H. Aoyama, S. Fujii, H. Sunagawa, H. Nagahama, M. Akutsu, N. Shinzato, N. Kaneko, & T. Nakamori. 2016. A quantitative protocol for DNA metabarcoding of springtails (Collembola). Genome 59:707-723.

Savory, T. 1964. Arachnida. Academic Press. New York, New York.

Shen, Y.Y., X. Chen, R.W. Murphy. 2013. Assessing DNA barcoding as a tool for species identification and data quality control. PLoS ONE 8:e57125.

Shendure, J., J. Hanlee. 2008. Next-generation DNA sequencing. Nature Biotechnology 26:1135-1145.

Sheppard, S.K., J. Bell, K.D. Sunderland, J. Fenlon, D. Skervin, & W.O.C. Symondson. 2005. Detection of secondary predation by PCR analyses of the gut contents of invertebrate generalist predators. Molecular Ecology 14:4461-4468.

Shaw, J.L., L.J. Clarke, S.D. Wedderburn, T.C. Barnes, L. S. Weyrich, & A. Cooper. 2016. Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. Biological Conservation 197:131-138.

Smrž, J., Ĺ. Kováč, J. Mikeš, & A. Lukešová. 2013. Microwhip scorpions (Palpigradi) feed on heterotrophic cyanobacteria in Slovak caves – a curiosity among Arachnida. PLoS ONE 8:e75989.

Spouge, J.L., & L. Mariño-Rameríz. 2012. The practical evaluation of DNA barcode efficacy. Methods in Molecular Biology 858:365-377.

Staats, M., A.J. Arulandhu, B. Gravendeel, A. Hols-Jensen, I. Scholtens, T. Peelen, T.W. Prins, & E. Kok. Advances in DNA metabarcoding for food and wildlife forensic species identification. Analytical and Bioanalytical Chemistry 408:4615-4630.

Symondson, W.O.C. 2002. Molecular identification of prey in predator diets. Molecular Ecology 11:627-641.

Thomsen, P.F., & E. Willerslev. 2015. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183:4-18.

Vestheim, H., & S.N. Jarmon. 2008. Blocking primers to enhance PCR amplification of rare sequences in mixed samples – a case study on prey DNA in Antarctic krill stomachs. Frontiers in Zoology 5:12.1-12.11.

Wilson, J.J. 2012. DNA Barcodes for Insects, Chapter 3 of DNA Barcodes: Methods and Protocols. Methods in Molecular Biology 858:17-46.

White, T.J., T.D. Bruns, S.B. Lee, & J.W. Taylor.1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and applications. United States: Academic Press. pp. 315–322.

Xu, C.C.Y., I.J. Yen, D. Bowman, & C.R. Turner. 2015. Spider Web DNA: A new spin on noninvasive genetics of predator and prey. PLoS ONE 10:e0142503.

Yang, C., X. Wang, J.A. Miller, M. de Blécourt, Y. Ji, C. Yang, R.D. Harrison, & D.W. Yu. 2014. Using metabarcoding to ask if easily collected soil and leaf-litter samples can be used as a general biodiversity indicator. Ecological Indicators 46:379-389.

Yilmaz, N., C.M. Visagie, J. Houbraken, J.C. Frisvad, & R.A. Samson. 2014. Polyphasic taxonomy of the genus. Studies in Mycology 78:175-341.

Zalar, P., G.S. De Hoog, H.J. Schroers, P.W. Crous, J.Z. Groenewald, & N. Gunde-Cimerman. 2007. Phylogeny and ecology of the ubiquitous saprobe *Cladosporium sphaerospermum*, with descriptions of seven new species from hypersaline environments. Studies in Mycology 58:157–183.

Zeale, M.R., R.K. Butlin, G.L.A. Barker, D.C. Lees, & G. Jones. 2011. Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. Molecular Ecology 11:236-244.

APPENDIX 1

PCR conditions for final dataset.-For PCR amplification of the ITS region, a 25 ul reaction was used consisting of: 5ul of 5X Phusion Buffer (New England Biolabs, Inc., Ipswich, MA), 0.15ul of 50mM MgCl, 1ul of the forward primer ITS3F(10uM), 1ul of the reverse primer ITS4R (10uM), 0.5ul of dNTPs (10mM each), 0.25ul of Phusion Polymerase – 2U/ul (New England Biolabs, Inc., Ipswich, MA), 3.7ul of DNA template (2ng), and 13.4ul of water. The thermal profile was 95˚C for 2 minutes, followed by 95˚C 30 seconds, 55˚C 30 sec, 72˚C 1 minute, for 5 cycles. Then 35 cycles of 95˚C, 30 seconds, 57.4˚C 30 seconds, 72˚ 1 minute, 72˚C 10 minute, and a 4˚C hold.

PCR amplification of the 16S rRNA gene was carried out in 20 ul reactions. consisting of: 4ul of 5X Phusion Buffer (New England Biolabs, Inc., Ipswich, MA), 0.2ul of 50mM $MgCl_2$, 1ul of the forward primer Cyan359F (10uM), 1ul of the reverse primer Cyan781R (10uM), 0.5ul of dNTPs (10mM each), 2 ul of BSA (20mg/ml), 0.2 ul of Phusion Polymerase – 2U/ul (New England Biolabs, Inc., Ipswich, MA), 3.7ul of DNA template (2ng), and 7.4ul of water. The thermal profile was 98˚C for 5 min, 94˚C 1 min, 62˚C-57˚C 1 min (increasing 0.5˚C per cycle) for 10 cycles, 72˚C 1 min, then 30 cycles of 94˚C 1 min, 57˚C 1 min 72˚C 1 min, 72˚C for 5 min.

For PCR amplification of the COI gene, reactions were carried out in 20 ul reactions consisting of: 4ul of 5X Phusion Buffer (New England Biolabs, Inc., Ipswich, MA), 1ul of the forward primer LepF1 (10uM), 1ul of the reverse primer MLepF1-Rev (10uM), 1 ul of the blocking primer (100mM), 0.5ul of dNTPs (10mM each), 2 ul of BSA (20mg/ml), 0.2 ul of Phusion Polymerase – 2U/ul (New England Biolabs, Inc., Ipswich, MA), 6ul of DNA template (2ng), and 4.3ul of water. The thermal profile was 98˚C for 1 min. then 94˚C 30 sec,

45˚C 40 sec, 72˚C 1 min. for 5 cycles, then 30-35 cycles of 94˚C 30 sec, 51˚C 40 sec, 72˚C 1 min, then 72˚C for 10 minutes (Wilson 2012).